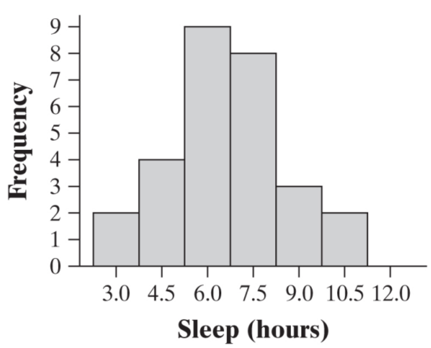
Section 9.3

Check Your Understanding, page 579:

1. H_0: μ = 320 versus H_a: $\mu \neq 320$, where μ = the true mean amount of active ingredient (mg) in Aspro tablets from this batch of production.
2. Random: The data come from a random sample. 10%: The sample of size 36 is less than 10% of the population of all tablets in this batch. Normal/Large Sample: $n = 36 \geq 30$. All conditions are met.
3. $t = \dfrac{319 - 320}{3/\sqrt{36}} = -2$.
4. For this test, df = 36 – 1 = 35. Using Table B and df = 30, the tail area is between 0.025 and 0.05. Thus, the *P*-value for the two-sided test is between 0.05 and 0.10. *Using technology*: 2tcdf(lower: –1000, upper: –2, df: 35) = 2(0.0267) = 0.0534. Because the *P*-value of 0.0534 is greater than α = 0.05, we fail to reject H_0. There is not convincing evidence that the true mean amount of the active ingredient in Aspro tablets from this batch of production differs from 320 mg.

Check Your Understanding, page 583:

1. *State*: We want to perform a test of $H_0: \mu = 8$ versus $H_a: \mu < 8$ where μ is the true mean amount of sleep that students at the professor's school get each night. We will perform the test at the $\alpha = 0.05$ significance level. *Plan*: If conditions are met, we should do a one-sample *t* test for the population mean μ. Random: The data come from a random sample. 10%: The sample size (28) is less than 10% of the population of students at this school. Normal/Large Sample: There were only 28 students so we need to examine the sample data. The histogram below indicates that there is not much skewness and no outliers, so it is reasonable to use a *t* procedure.



Do: The sample mean and standard deviation are: $\bar{x} = 6.643$ and $s_x = 1.981$. The corresponding test statistic is $t = \dfrac{6.643 - 8}{1.981/\sqrt{28}} = -3.625$. With df = 27, the *P*-value is between 0.0005 and 0.001. *Using technology*: *P*-value = 0.0006. *Conclude*: Because our *P*-value of 0.0006 is less than α = 0.05, we reject H_0. There is convincing evidence that students at this university get less than 8 hours of sleep, on average.

Check Your Understanding, page 586:

1. *State*: We want to perform a test of $H_0 : \mu = 128$ versus $H_a : \mu \neq 128$ where μ is the true mean systolic blood pressure for the company's middle-aged male employees. We will perform the test at the $\alpha = 0.05$ significance level. *Plan*: If conditions are met, we should do a one-sample *t* test for the population mean μ. Random: The data come from a random sample. 10%: The sample size (72) is less than 10% of the population of middle-aged male employees at this large company. Normal/Large Sample: $n = 72 \geq 30$. *Do*: From the computer output, $t = 1.10$ and *P*-value = 0.275 *Conclude*: Because our *P*-value of 0.275 is greater than $\alpha = 0.05$, we fail to reject H_0. There is not convincing evidence that the mean systolic blood pressure for this company's middle-aged male employees differs from the national average of 128.
2. We are 95% confident that the interval from 126.43 to 133.43 captures the true mean systolic blood pressure for the company's middle-aged male employees. The value of 128 is in this interval and therefore is a plausible mean systolic blood pressure for the males 35 to 44 years of age. This agrees with the results of the significance test, which did not reject a mean value of 128.

Check Your Understanding, page 589:

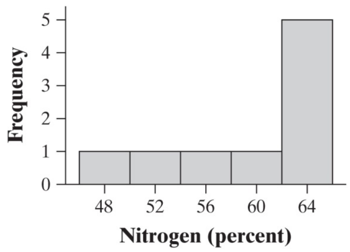
State: We want to perform a test of $H_0 : \mu_d = 0$ seconds versus $H_a : \mu_d > 0$ seconds where μ_d is the true mean difference (air – nitrogen) in pressure lost. We will perform the test at the $\alpha = 0.05$ significance level. *Plan*: If conditions are met, we should do a paired *t* test for μ_d. Random: Treatments were assigned at random to each pair of tires. Normal/Large Sample: The number of differences is large ($n = 31 \geq 30$). *Do*: The sample mean and standard deviation are: $\bar{x} = 1.252$ and $s_x = 1.202$. The corresponding test statistic is $t = \frac{1.252 - 0}{1.202/\sqrt{31}} = 5.80$ With df = 30, *P*-value < 0.0005. *Using technology*: *P*-value $\approx$ 0. *Conclude*: Because the *P*-value of approximately 0 is less than $\alpha = 0.05$, we reject H_0. We have convincing evidence that the true mean difference in pressure lost (air – nitrogen) is greater than 0. In other words, we have convincing evidence that tires lose less pressure when filled with nitrogen than when filled with air, on average.

Exercises, page 595:

9.65 Random: The sample was randomly selected. 10%: The sample size (45) is less than 10% of the population size of 1000. Normal/Large Sample: $n = 45 \geq 30$.

9.66 Random: The sample was randomly selected. 10%: The sample size (50) is less than 10% of the population of children in Jordan. Normal/Large Sample: $n = 50 \geq 30$.

9.67 The Random condition may not be met, because we don't know if this is a random sample of the atmosphere in the Cretaceous era. Also, the Normal/Large Sample condition is not met. The sample size is less than 30 and the histogram below shows that the data are skewed to the left.



9.68 The Normal/Large Sample condition is not met. The sample size is less than 30 and the histogram shows that the data are strongly skewed to the left. Also, because the lower boundary for outliers is 68 – 1.5(100 – 68) = 20, there are two low outliers.

9.69 (a) The test statistic is $t = \dfrac{125.7 - 115}{29.8/\sqrt{45}} = 2.409$.

(b) For this test, df = 45 – 1 = 44. Using Table B and df = 40, $0.01 < P\text{-value} < 0.02$. Using technology: P-value = tcdf(lower: 2.409, upper: 1000, df: 44) = 0.0101.

9.70 (a) The test statistic is $t = \dfrac{11.3 - 12}{1.6/\sqrt{50}} = -3.094$.

(b) For this test, df = 50 – 1 = 49. Using Table B and df = 40, $0.001 < P\text{-value} < 0.0025$. *Using technology*: P-value = tcdf(lower: –1000, upper: –3.094, df = 49) = 0.0016.

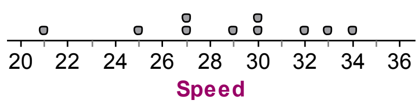
9.71 (a) Using Table B and df = 19, $0.025 < P\text{-value} < 0.05$. *Using technology*: P-value = tcdf(lower: –1000, upper: –1.81, df: 19) = 0.043. *5% significance level*: Because the P-value of 0.043 is less than α = 0.05, we reject H_0. There is convincing evidence that $\mu < 5$. *1% significance level*: Because the P-value of 0.043 is greater than α = 0.01, we fail to reject H_0. There is not convincing evidence that $\mu < 5$.

(b) If the alternative hypothesis is two-sided, $0.05 < P\text{-value} < 0.10$. *Using technology*: P-value = 2tcdf(lower: –1000, upper: –1.81, df: 19) = 2(0.043) = 0.086. *5% significance level*: Because the P-value of 0.086 is greater than α = 0.05, we fail to reject H_0. There is not convincing evidence that $\mu \neq 5$. *1% significance level*: Because the P-value of 0.086 is greater than α = 0.01, we fail to reject H_0. There is not convincing evidence that $\mu \neq 5$.

9.72 (a) Using Table B and df = 24, $2(0.10) < P\text{-value} < 2(0.15)$ or $0.20 < P\text{-value} < 0.30$. *Using technology*: P-value = 2tcdf(lower: –1000, upper: –1.12, df = 24) = 0.2738. *5% significance level*: Because the P-value of 0.2738 is greater than α = 0.05, we fail to reject H_0. There is not convincing evidence that $\mu \neq 64$. *1% significance level*: Because the P-value of 0.2738 is greater than α = 0.01, we fail to reject H_0. There is not convincing evidence that $\mu \neq 64$.

(b) *Using technology*: *P*-value = tcdf(lower: –1000, upper: –1.12, df = 24) = 0.1369. *5% significance level*: Because the *P*-value of 0.1369 is greater than α = 0.05, we fail to reject H_0. There is not convincing evidence that $\mu < 64$. *1% significance level*: Because the *P*-value of 0.1369 is greater than α = 0.01, we fail to reject H_0. There is not convincing evidence that $\mu <$ 64.

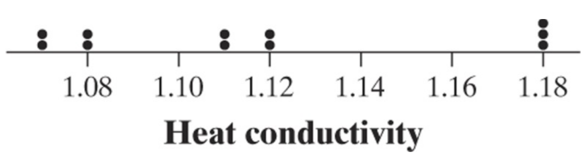
9.73 (a) *State*: We want to perform a test of H_0: μ = 25 versus H_a: μ > 25 where μ is the true mean speed of all drivers in a construction zone. We will use α = 0.05. *Plan*: If conditions are met, we should do a one-sample *t* test for the population mean μ. Random: The drivers were randomly selected. 10%: The sample size (10) is less than 10% of all drivers in a construction zone. Normal/Large Sample: Because the sample is small, we need to graph the sample data. There is no strong skewness or outliers in the sample, so it is reasonable to use a *t* procedure.



Do: The sample mean and standard deviation are $\bar{x}$ = 28.8 and s_x = 3.94. The test statistic is $t = \dfrac{28.8 - 25}{3.94/\sqrt{10}} = 3.05$. With df = 10 – 1 = 9, the *P*-value is between 0.005 and 0.01. *Using technology*: *P*-value = 0.0069. *Conclude*: Because the *P*-value of 0.0069 is less than α = 0.05, we reject H_0. We have convincing evidence that the true mean speed of all drivers in the construction zone is greater than 25 mph.

(b) Because we rejected H_0, it is possible we made a Type I error—finding convincing evidence that the true mean speed is greater than 25 mph when it really isn't.

9.74 (a) *State*: We want to perform a test of $H_0: \mu = 1$ versus $H_a: \mu > 1$ where μ is the true mean conductivity for this type of glass. We will perform the test at the $\alpha = 0.05$ significance level. *Plan*: If conditions are met, we should do a one-sample *t* test for the population mean μ. Random: The sample was randomly selected. 10%: The sample size (11) is less than 10% of all pieces of this particular type of glass. Normal/Large Counts: There were only 11 pieces sampled so we need to examine the sample data. The dotplot below indicates that there is not much skewness and no outliers, so it is reasonable to use a *t* procedure.



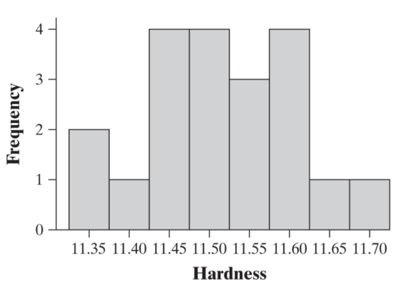
Do: The sample mean and standard deviation are: $\bar{x} = 1.1182$ and $s_x = 0.0438$. The corresponding test statistic is $t = \dfrac{1.1182 - 1}{0.0438/\sqrt{11}} = 8.95$. With df = 10, the *P*-value is < 0.0005. Using technology, *P*-value ≈ 0. *Conclude*: Because the *P*-value of approximately 0 is less than α = 0.05, we reject H_0. There is convincing evidence that this glass has a true mean conductivity greater than 1.

(b) Because we rejected H_0, it is possible we made a Type I error—finding convincing evidence that the true mean conductivity is greater than 1 when it really isn't.

9.75 (a) *State*: We want to perform a test of $H_0: \mu = 1200$ versus $H_a: \mu < 1200$ where μ is the true mean daily calcium intake of women 18 to 24 years of age. We will perform the test at the $\alpha = 0.05$ significance level. *Plan*: If conditions are met, we should do a one-sample *t* test for the population mean μ. Random: The sample was randomly selected. 10%: The sample size (36) is less than 10% of all women aged 18 to 24. Normal/Large Sample: $n = 36 \geq 30$. *Do*: From the computer output, $t = -6.73$ and *P*-value = 0.000. *Conclude*: Because the *P*-value of approximately 0 is less than $\alpha = 0.05$, we reject H_0. There is convincing evidence that women aged 18 to 24 are getting less than 1200 mg of calcium daily, on average.
(b) Assuming that women aged 18 to 24 get 1200 mg of calcium per day, on average, there is about a 0 probability that we would observe a sample mean less than or equal to 856.2 mg by chance alone.

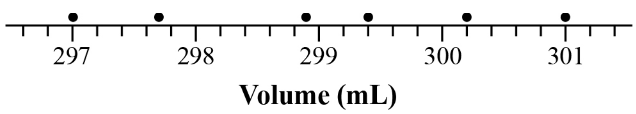
9.76 (a) *State*: We want to perform a test of $H_0: \mu = 0.95$ versus $H_a: \mu < 0.95$ where μ is the true mean weekly return on the broker's portfolios. We will perform the test at the $\alpha = 0.05$ significance level. *Plan*: If conditions are met, we should do a one-sample *t* test for the population mean μ. Random: The sample was randomly selected. 10%: The sample size (36) is less than 10% of the total number of weeks in the 10 years (520). Normal/Large Sample: $n = 36 \geq 30$. *Do*: From the computer output, $t = -2.98$ and *P*-value = 0.003. *Conclude*: Because our *P*-value of 0.003 is less than $\alpha = 0.05$, we reject H_0. We have convincing evidence that the portfolios managed by this broker had a smaller true mean return than the Standard and Poor's mean return.
(b) Assuming that the broker's true mean weekly return is 0.95%, there is a 0.003 probability of getting a sample mean weekly return of –1.441% or less by chance alone.

9.77 *State*: We want to perform a test of $H_0: \mu = 11.5$ versus $H_a: \mu \neq 11.5$ where μ is the true mean hardness of the tablets. We will use $\alpha = 0.05$. *Plan*: If conditions are met, we should do a one-sample *t* test for the population mean μ. Random: The tablets were selected randomly. 10%: The sample size (20) is less than 10% of all tablets in the batch. Normal/Large Sample: Because the sample is small, we need to graph the sample data. There is no strong skewness or outliers in the sample, so it is reasonable to use a *t* procedure.



Do: The sample mean and standard deviation are $\bar{x}$ = 11.5164 and s_x = 0.0950. The test statistic is $t = \dfrac{11.5164 - 11.5}{0.0950/\sqrt{20}} = 0.77$. With df= 20 – 1 = 19, the area in one tail is between 0.20 and 0.25. Thus, the *P*-value is between 0.40 and 0.50. *Using technology*: *P*-value = 0.4494. *Conclude*: Because our *P*-value of 0.4494 is greater than α = 0.05, we fail to reject H_0. We do not have convincing evidence that the true mean hardness of these tablets is different from 11.5.

9.78 *State*: We want to perform a test of $H_0: \mu = 300$ml versus $H_a: \mu \neq 300$ml where μ is the true mean amount of cola in the bottles. We will perform the test at the $\alpha = 0.05$ significance level. *Plan*: If conditions are met, we should do a one-sample *t* test for the population mean μ. Random: The bottles were selected randomly. 10%: The sample size (6) is less than 10% of all bottles produced this day. Normal/Large Sample: Because the sample is small, we need to graph the sample data. There is no strong skewness or outliers in the sample, so it is reasonable to use a *t* procedure.



Do: The sample mean and standard deviation are: $\bar{x} = 299.033$ and $s_x = 1.503$. The test statistic is $t = \dfrac{299.033 - 300}{1.503/\sqrt{6}} = -1.58$. With df= 6 – 1 = 5, the area in one tail is between 0.05 and 0.10. Thus, the *P*-value is between 0.10 and 0.20. *Using technology*: *P*-value = 0.1760. *Conclude*: Because our *P*-value of 0.1760 is greater than α = 0.05, we fail to reject H_0. We do not have convincing evidence that the true mean amount of cola in all the bottles differs from the target value of 300 ml.

9.79 *State*: We want to estimate μ = the true mean hardness for this type of pill at a 95% confidence level. *Plan*: We should construct a one-sample *t* interval for μ if the conditions are met. The conditions are all met (see Exercise 9.77). *Do*: From the data we find that $\bar{x} = 11.516$, $s_x = 0.095$ and $n = 20$. With df = 19, $t^* = 2.093$. The confidence interval is:
$11.516 \pm 2.093\left(\dfrac{0.095}{\sqrt{20}}\right) = 11.516 \pm 0.044 = (11.472, 11.561)$. *Conclude*: We are 95% confident that the interval from 11.472 to 11.561 captures the true mean hardness measurement for this type of pill. The confidence interval agrees with the test done in Exercise 77. Both give 11.5 as a plausible value for the true mean hardness μ. The confidence interval, however, gives other plausible values as well.

9.80 *State*: We want to estimate μ = the true mean amount of cola in the bottles at a 95% confidence level. *Plan*: We should construct a one-sample *t* interval for μ if the conditions are met. The conditions are all met (see Exercise 9.78). *Do*: From the data we find that $\bar{x}$ = 299.033, s_x = 1.503, and n = 6. With df = 5, $t^* = 2.571$. The confidence interval is:

$299.033 \pm 2.571\left(\frac{1.503}{\sqrt{6}}\right) = 299.033 \pm 1.578 = (297.455, 300.611)$. *Conclude*: We are 95% confident that the interval from 297.455 to 300.611 ml captures the true mean amount of cola in the bottles. The confidence interval agrees with the test done in Exercise 78. Both give 300 ml as a plausible value for the true mean amount of cola μ. The confidence interval, however, gives other plausible values as well.

9.81 *State*: We want to perform a test of $H_0: \mu = 200$ milliseconds versus $H_a: \mu \neq 200$ milliseconds where μ is the true mean response time of European servers. We will perform the test at the $\alpha = 0.05$ significance level. *Plan*: If conditions are met, we should do a one-sample *t* interval to help us perform a two-sided test for the population mean μ. Random: The servers were selected randomly. 10%: The sample size (14) is less than 10% of all servers in Europe. Normal/Large Sample: The sample size is small, but a graph of the data reveals no strong skewness or outliers, so it is reasonable to use a *t* procedure. *Do*: The 95% confidence interval from software is (158.22, 189.64). *Conclude*: Because our 95% confidence interval does not contain 200 milliseconds, we reject H_0 at the $\alpha = 0.05$ significance level. We have convincing evidence that the mean response time of European servers is different from 200 milliseconds.

9.82 *State*: We want to perform a test of $H_0: \mu = 5$ versus $H_a: \mu \neq 5$ where μ is the true mean number of glasses of water that a U.S. adult drinks per day. We will perform the test at the $\alpha = 0.10$ significance level. *Plan*: If conditions are met, we should do a one-sample *t* interval to help us perform a two-sided test for the population mean μ. Random: The adults were selected randomly. 10%: The sample size (24) is less than 10% of the population of U.S. adults. Normal/Large Sample: The sample size is small, but a graph of the data reveals a roughly symmetric shape with no outliers, so it is reasonable to use a *t* procedure. *Do*: The 90% confidence interval from software is (3.794, 4.615). *Conclude*: Because our 90% confidence interval does not contain 5 glasses of water, we reject H_0. We have convincing evidence that the true mean number of glasses of water that U.S. adults drink in a day is different from 5.

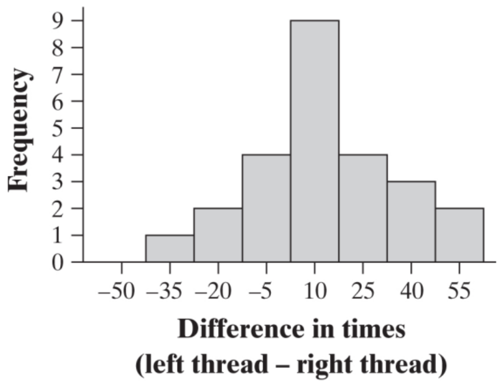
9.83 (a) Yes. Because the *P*-value of 0.06 is greater than α = 0.05, we fail to reject H_0: μ = 10 at the 5% level of significance. Thus, the 95% confidence interval will include 10.
(b) No. Because the *P*-value of 0.06 is less than α = 0.10, we reject $H_0: \mu = 10$ at the 10% level of significance. Thus, the 90% confidence interval would not include 10 as a plausible value.

9.84 (a) Yes. Because the *P*-value of 0.03 is greater than α = 0.01, we fail to reject H_0: μ = 15 at the 1% level of significance. Thus, the 99% confidence interval will include 15.
(b) No. Because the *P*-value of 0.03 is less than α = 0.05, we reject $H_0: \mu = 15$ at the 5% level of significance. Thus, the 95% confidence interval would not include 15 as a plausible value.

9.85 (a) Randomizing the order balances out the effects of other variables that might affect the times for each thread. For example, if all the subjects used the right thread first and they were

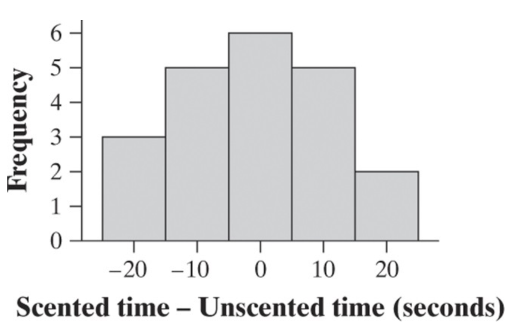
tired when they used the left thread, then we wouldn't know if the difference in times was because of tiredness or because of the direction of the thread.

(b) *State:* We want to perform a test of $H_0 : \mu_d = 0$ versus $H_a : \mu_d > 0$ where μ_d is the true mean difference (left – right) in the time (in seconds) it takes to turn the knob with the left-hand thread and the right-hand thread. We will use $\alpha = 0.05$. *Plan*: If conditions are met, we should do a paired *t* test for μ_d. Random: The order of treatments was determined at random. Normal/Large Sample: Because the number of differences is small, we need to graph the differences. There is no strong skewness or outliers, so it is reasonable to use a *t* procedure.



Do: The sample mean and standard deviation are: $\bar{x} = 13.32$ and $s_x = 22.94$. The test statistic is $t = \dfrac{13.32 - 0}{22.94/\sqrt{25}} = 2.903$. With df = 25 – 1 = 24, the *P*-value is between 0.0025 and 0.005. *Using technology*: *P*-value = 0.0039. *Conclude*: Because the *P*-value of 0.0039 is less than $\alpha = 0.05$, we reject H_0. We have convincing evidence that the true mean difference (left – right) in time it takes to turn the knob is greater than 0. That is, we have convincing evidence that right-handed people find right-handed threads easier to use.

9.86 (a) Randomizing the order balances out the effects of other variables that might affect the times for each scent. For example, if all the subjects used the scented mask first and they were tired when they used the unscented mask, then we wouldn't know if the difference in times was because of tiredness or because of the mask they were using.
(b) *State:* We want to perform a test of $H_0 : \mu_d = 0$ seconds versus $H_a : \mu_d < 0$ seconds where μ_d is the true mean difference (scented – unscented) in time it takes to complete the mazes with the scented and unscented masks. We will perform the test at the $\alpha = 0.05$ significance level. *Plan*: If conditions are met, we should do a paired *t* test for μ_d. Random: The order of treatments was determined at random. Normal/Large Sample: Because the number of differences is small, we need to graph the differences. There is no strong skewness or outliers, so it is reasonable to use a *t* procedure.



Do: The sample mean and standard deviation are: $\bar{x} = -0.96$ and $s_x = 12.55$. The corresponding test statistic is $t = \dfrac{-0.96 - 0}{12.55/\sqrt{21}} = -0.351$. With df = 20, *P*-value > 0.25. *Using technology*: *P*-value = 0.3646. *Conclude*: Because the *P*-value of 0.3646 is greater than α = 0.05, we fail to reject H_0. We do not have convincing evidence that the true mean difference in time (scented – unscented) is less than 0. In other words, we don't have convincing evidence that it takes people less time to complete the maze with the scented mask than with the unscented mask, on average.

9.87 (a) $H_0 : \mu_d = 0$ versus $H_a : \mu_d > 0$ where μ_d is the true mean difference in tomato yield (A – B).
(b) df = number of differences – 1 = 10 – 1 = 9.
(c) *Interpretation*: Assuming that the average yield for both varieties is the same, there is a 0.1138 probability of getting a mean difference as large or larger than the one observed in this experiment. *Conclusion*: Because the *P*-value of 0.1138 is greater than α = 0.05, we fail to reject H_0. We do not have convincing evidence that the true mean difference in tomato yield (A – B) is greater than 0. In other words, we do not have convincing evidence that Variety A tomato plants have a greater mean yield.
(d) A Type I error is finding convincing evidence that Variety A tomato plants have a greater mean yield, when in reality there is no difference. A Type II error is not finding convincing evidence that Variety A tomato plants have a higher mean yield, when in reality Variety A does have a greater mean yield. Because the researchers didn't find convincing evidence that Variety A tomato plants have a higher mean yield, they might have made a Type II error.

9.88 (a) $H_0 : \mu_d = 0$ versus $H_a : \mu_d < 0$ where μ_d is the true mean difference in number of words recalled (music – silence).
(b) df = number of differences – 1 = 30 – 1 = 29.
(c) *Interpretation*: Assuming that the average number of words recalled is the same when listening to music and in silence, there is a 0.0027 probability of getting a mean difference as small or even smaller than the mean difference observed in our experiment. *Conclusion*: Because the *P*-value of 0.0027 is less than α = 0.05, we reject H_0. We have convincing evidence that the true mean difference in number of words recalled is less than 0. In other words, we have convincing evidence that listening to music hinders students learning.
(d) A Type I error is finding convincing evidence that the average number of words recalled when listening to music is less than when in silence, when in reality there is no difference. A Type II error is not finding convincing evidence that the average number of words recalled when listening to music is less than when in silence, when in reality the average number of words recalled is smaller when listening to music. Because the researchers did find convincing evidence that the average is smaller when listening to music, they might have made a Type I error.

9.89 To increase power, the researchers could increase the significance level α or increase the sample size *n*.

9.90 Using $\alpha = 0.10$ would increase the power of the test. Increasing α means increasing the probability of a Type I error. This reduces the probability of a Type II error and consequently increases power. A higher significance level also makes it easier to reject H_0, leading to more correct rejections of H_0 when it is false and, thus, greater power.

9.91 When the sample size is very large, rejecting the null hypothesis is very likely, even if the actual parameter is only slightly different from the hypothesized value. However, there may be no practical importance in a difference this small.

9.92 This sample wasn't randomly selected—it was a convenience sample. Depending on the time of day or the day of the week, certain types of shoppers may be underrepresented. If these shoppers spend more (or less) money on average than the population of shoppers at the store, the estimated mean will be too low (or too high).

9.93 (a) No, in a sample of size $n = 500$, we expect to see about (500)(0.01) = 5 people who do better than random guessing, with a significance level of 0.01. These four might have ESP, or they may simply be among the "lucky" ones we expect to see just by chance.
(b) The researcher should repeat the procedure on these four to see if they again perform well.

9.94 This is not information taken from a sample. We have information about all presidents—the whole population of interest—so Joe can calculate the exact value of μ.

9.95 b

9.96 a

9.97 d

9.98 c

9.99 c

9.100 a

9.101 a

9.102 b

9.103 (a) Not included. The margin of error does not account for undercoverage.
(b) Not included. The margin of error does not account for nonresponse.
(c) Included. The margin of error is calculated to account for sampling variability.

9.104 (a) Note that $P(\text{at least one apple}) = 1 - P(\text{no apples})$. We are told that the 5 wheels spin independently and that $P(\text{apple}) = 0.20$. So
$P(\text{at least one apple}) = 1 - P(\text{no apples}) = 1 - (0.80)^5 = 0.67232$. There is a 0.67232 probability of getting at least one apple in one pull of the lever.
(b) To answer the reader's question we assume that all pulls of the lever are independent. We know that the probability of no apples on a given pull of the lever is $1 - 0.67232 = 0.32768$. So

the probability of no apples on any of 5 pulls is $(0.32768)^5 = 0.00378$ and the probability of at least one apple in 5 pulls of the lever is $1 - 0.00378 = 0.99622$.